

Mu2e Data Management Plan

31 August 2015

Doug Glenzinski, Rob Kutschke, Jim Miller

Experiment Description

The Mu2e experiment, located at Fermilab, will search for the neutrino-less conversion of a muon into an electron while in the field of a nucleus, $\mu N \rightarrow e N$. This conversion process is an example of charged-lepton flavor violation, which has never been observed but which is predicted to occur in a wide array of new physics models. Mu2e will achieve a sensitivity that is 10,000 times better than the current world's best search for muon-to-electron conversion [1] and will probe new physics mass scales up to 10,000 TeV/c² [2]. The muons required for the experiment are produced by colliding 8 GeV protons onto a tungsten production target. The resulting pions are collected and allowed to decay (predominantly to muons). The muons are stopped on an aluminum stopping target. Active detector elements are located downstream and measure the position, timing, momentum, and energy of particles emanating from the stopping target. The production target, pion transport and decay pipe, stopping target, and detectors are all immersed in a magnetic field produced by a system of three custom-built solenoids. Mu2e is scheduled to begin commissioning in 2020 and will require a 4 year run once the beam and the detector are ready for physics. It is estimated that Mu2e will produce a few PB of data each year.

DOE's Roles in the Experiment

The Mu2e experiment is being constructed as a DOE 413.3B project and received CD-2 (baseline) approval in March 2015. Final construction approval, CD-3c, is scheduled for summer of 2016. Fermilab is the lead laboratory for the Mu2e project and is responsible for delivering the full scope of the project with the support of the Mu2e collaboration. It is planned that a significant fraction of the Mu2e calorimeter will be delivered in-kind by the Italian Istituto Nazionale di Fisica Nucleare (INFN). Other international partners include Russia and Germany, who are primarily collaborating with the INFN on the calorimeter. Fermilab is the host institution for the Mu2e Collaboration and together they are ultimately responsible for the operations and management of the Mu2e data, beam line, and detector.

Organization – Agency/Lab Level

The Mu2e group in the Muon Department of Fermilab's Particle Physics Division (PPD) is home to Fermilab scientists whose primary research effort is on the Mu2e experiment. The head of the Mu2e group together with the collaboration co-spokespersons serve as the primary liaisons between the Mu2e collaboration and the laboratory. Fermilab's PPD is the lead division responsible for ensuring the safe and efficient operation of the Mu2e apparatus. The Fermilab Accelerator Division (AD) is responsible for safely delivering beam to Mu2e while the Scientific Computing Division (SCD) is responsible for providing the infrastructure and expert support required to manage and store Mu2e data and software. Fermilab's other Divisions and Sections provide support for the safe and efficient operation of Mu2e as needed.

Organization – Experiment Level

During the experiment's design and construction phase the Mu2e project and collaboration work together to deliver the project scope. The organization of and relationship between the project and collaboration are represented by the chart below. The project is organized into nine L2 sub-systems and is led by the project manager, who takes input from a Technical Board, a Risk Management Board, and a Change Control Board. The collaboration is organized into working groups and is led by two equal co-spokespersons, who are advised by an Executive Board and an Institutional Board. Once the project is completed and detector operations begin, the collaboration works with Fermilab to ensure safe and efficient data collection, processing, and physics analysis.

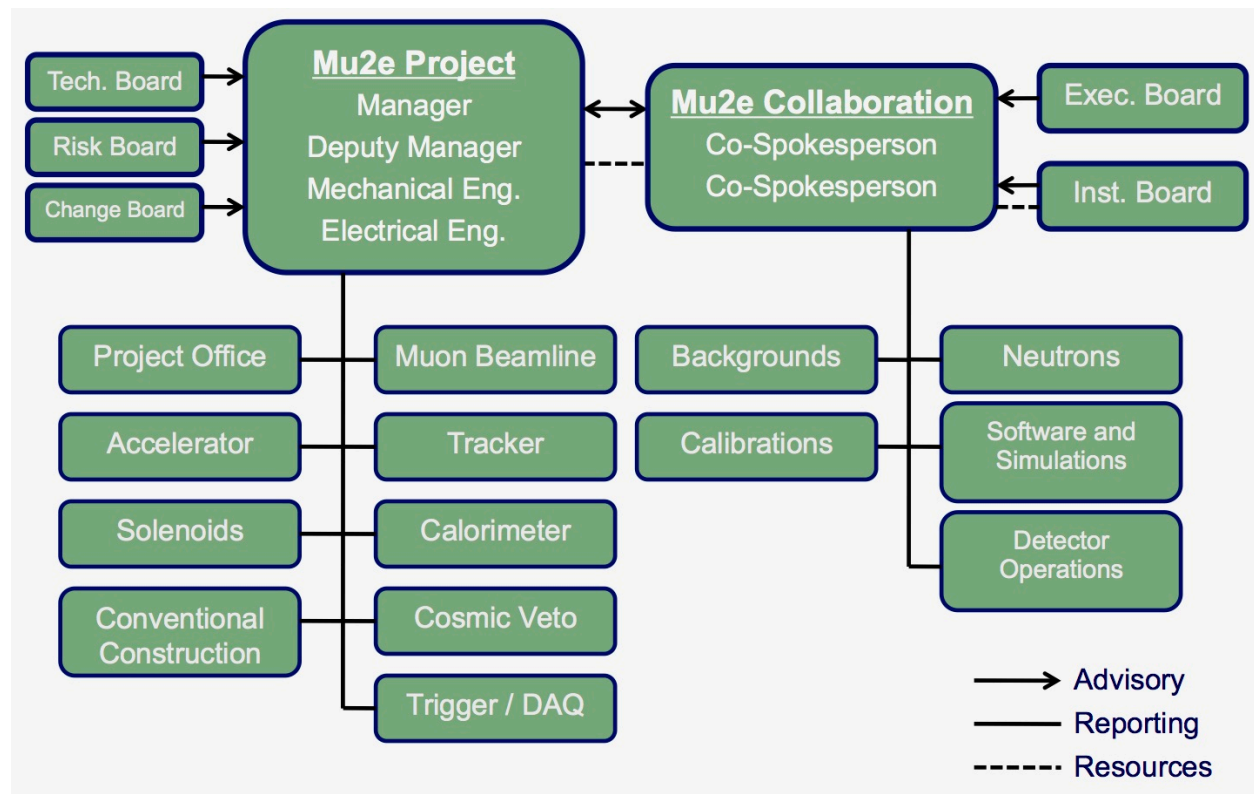


Figure 1: Mu2e organization chart for the project and collaboration.

Collaboration

The Mu2e collaboration is comprised of approximately 170 scientists from 33 institutions spread among four countries. The collaboration is led by two equal co-spokespersons elected from the collaboration membership for staggered two-year terms.

Data Policy Management

With direction from the collaboration, the head of the Mu2e Software and Simulations working group establishes the protocols and policies affecting Mu2e data, simulations, and software. The policies and protocols are developed in consultation with the Fermilab Scientific Computing Division (SCD) and wherever possible utilize infrastructure, protocols, and policies provided by and managed by SCD. The resources provided by SCD are summarized in reference [3]. To facilitate clear communication and close cooperation SCD provides a scientist liaison to work with Mu2e. A Technical Statement of Work formalizes in detail the roles and responsibilities of the Mu2e collaboration and SCD.

Data Description and Processing

The data will be recorded in Raw format. After calibration and reconstruction the data will again be saved in a Reconstructed format. After additional high level processing and event skimming, Analysis datasets will be formed. The Raw, Reconstructed, and Analysis data formats will exist for both experimental data and simulated events.

All formats of all data will be stored at Fermilab using the SCD provided facilities. The Raw data from the experiment will be written to tape storage in real time with a maximum delay of about 1 day. Raw data will be processed to produce Reconstructed data with a latency of 4-6 weeks to allow for finalizing and validating the calibration constants. Once established, the Analysis data sets will be produced in conjunction with the Reconstructed data.

Simulated events are written to tape storage with a delay of about 1 week to allow for validation. Most Mu2e simulations employ a staged approach in order to more efficiently use and manage the available CPU and storage resources. The event-data for each stage is stored separately as each stage is completed. Once the full set of stages are done, Reconstructed and, if applicable, Analysis data sets are produced and stored.

It is anticipated that roughly once each year all Raw data will be reprocessed to take advantage of improved reconstruction algorithms and calibration updates. During commissioning and the early stages of operations, reprocessing will be done several times per year. The resulting Reconstructed and Analysis data will be stored.

Data Products and Releases

Data is released to the collaboration as each data set is completed and after some standard validation has been performed. Available data sets are documented in summary form in the Mu2e document database and in detailed form in the SAM database provided by SCD for data handling purposes.

Software is developed in the environment provided by the *art* framework, developed and maintained by Fermilab's SCD. Mu2e specific code is managed using the git source code

management system. Frozen releases are established by the head of the Mu2e Software and Simulations working group. The releases are announced to the collaboration via email and recorded in the Mu2e document database. The code is stored on central servers that are backed-up daily by SCD.

Mu2e uses the standard e-log provided by SCD, which is backed-up daily by SCD.

Databases important for the operation of Mu2e or for the reconstruction and analysis of Mu2e data use standard facilities provided by SCD. These databases are backed-up daily by SCD.

Plan for Serving Data to the Collaboration and Community

Mu2e data are freely available to the entire Mu2e collaboration via the data handling services provided by Fermilab's SCD.

Plan for Archiving Data

All data will be digitally archived at Fermilab for several years after the experiment ends. The exact duration is yet to be determined. Mu2e will follow the recommended SCD policy for migration of tape based datasets to new generation media [3]. Mu2e will work with SCD to ensure that the Mu2e software can be built and will run correctly on the archived datasets so long as those datasets are retained.

Plan for Making Data Used in Publications Available

All data points shown in published graphs and tables will be also available in a machine readable form either on a website listed in the publication or using facilities provided by INSPIRE [4] and the Durham HepData project [5].

Responsiveness to SC Statement on Digital Data Management

This data management plan fully follows the Office of Science (SC) Statement on Digital Data Management except that not all Mu2e data is planned to be publicly available due to resource limitations.

References

- [1] W. Bertl *et al.*, Eur. Phys. J. **C47**, 337 (2006).
- [2] L. Bartoszek *et al.*, Mu2e Technical Design Report, Fermilab-TM-2594 (2014); available from <http://mu2e-docdb.fnal.gov/cgi-bin/ShowDocument?docid=4299>
- [3]http://computing.fnal.gov/xms/Science_&_Computing/Policies_and_Publications/Data_Management_Practices_and_Policies
- [4] <http://inspirehep.net>
- [5] <http://hepdata.cedar.ac.uk>